



THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

### Structure of the human VIPR2 gene for vasoactive intestinal peptide receptor type 2

**Citation for published version:**

Lutz, EM, Shen, S, Mackay, M, West, K & Harmar, AJ 1999, 'Structure of the human VIPR2 gene for vasoactive intestinal peptide receptor type 2', *FEBS Letters*, vol. 458, no. 2, pp. 197-203.  
[https://doi.org/10.1016/S0014-5793\(99\)01135-7](https://doi.org/10.1016/S0014-5793(99)01135-7)

**Digital Object Identifier (DOI):**

[10.1016/S0014-5793\(99\)01135-7](https://doi.org/10.1016/S0014-5793(99)01135-7)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Publisher's PDF, also known as Version of record

**Published In:**

FEBS Letters

**Publisher Rights Statement:**

Copyright 1999 Federation of European Biochemical Societies

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# Structure of the human VIPR2 gene for vasoactive intestinal peptide receptor type 2

E.M. Lutz<sup>a,\*</sup>, S. Shen<sup>b</sup>, M. Mackay<sup>b,1</sup>, K. West<sup>b</sup>, A.J. Harmar<sup>b</sup>

<sup>a</sup>MRC Brain Metabolism Unit, University Department of Neuroscience, 1 George Square, Edinburgh EH8 9JZ, UK

<sup>b</sup>MRC Brain Metabolism Unit, Royal Edinburgh Hospital, Morningside Park, Edinburgh EH10 5HF, UK

Received 22 June 1999

**Abstract** The VPAC<sub>2</sub> (vasoactive intestinal peptide (VIP)<sub>2</sub>) receptor is a seven-transmembrane spanning G protein-coupled receptor which responds similarly to VIP and pituitary adenylate cyclase activating polypeptide (PACAP) in stimulating cAMP production. Recently, we reported the localisation of the human VPAC<sub>2</sub> receptor gene (VIPR2) to chromosome 7q36.3 (Mackay, M. et al. (1996) *Genomics* 37, 345–353). Here, we describe the characterisation of the VIPR2 gene structure and promoter region. The VIPR2 gene is encoded by 13 exons, the initiator codon of the 438 amino acid open reading frame is located in exon 1 and the termination signal and a poly-adenylation signal sequence are located in exon 13. The 5' untranslated region extends 187 bp upstream of the initiator codon and is extremely GC-rich (80%). The poly-adenylation signal is located 2416 bp downstream of the stop codon. Intron sizes range from 68 bp (intron 11) to 45 kb (intron 4) and the human gene spans 117 kb.

© 1999 Federation of European Biochemical Societies.

**Key words:** Vasoactive intestinal peptide receptor type 2; Vasoactive intestinal peptide; Pituitary adenylate cyclase activating polypeptide; Vasoactive intestinal peptide receptor type 2 gene; Human

## 1. Introduction

The 28 amino acid vasoactive intestinal peptide (VIP) and the 38 amino acid pituitary adenylate cyclase activating polypeptide (PACAP) belong to a family of structurally related hormones which include secretin, growth hormone releasing hormone (GHRH), glucagon and glucagon-like peptide. The action of these hormones is mediated through receptors which belong to the group II family of seven-transmembrane G protein-coupled receptors first defined by cloning of the rat secretin, porcine calcitonin and opossum parathyroid hormone (PTH) receptors [2]. Receptors which are members of this family couple to adenylyl cyclase stimulation. Two distinct receptors which respond equally to VIP and PACAP have been cloned, the VPAC<sub>1</sub> and the VPAC<sub>2</sub> receptors [3–5]. Only 50% of their amino acid sequences are conserved, with the greatest identity in the transmembrane spanning domains. Both receptors are similarly related to the secretin [6] and the PACAP (PAC<sub>1</sub>) [7,8] receptors. The VPAC<sub>1</sub> receptor is expressed in liver, intestine and in the brain in cortex, hippocampus, cerebellum and olfactory bulb [3,9]. The VPAC<sub>2</sub> re-

ceptor is expressed in the olfactory bulb, thalamus, hippocampus and hypothalamus, particularly in the supra-chiasmatic nucleus [4,9,10], as well as in the pancreas, heart, stomach, lung and digestive tract [9,11]. The VPAC<sub>1</sub> and VPAC<sub>2</sub> receptors have been mapped to human chromosomal regions 3p22 [12] and 7q36.3 [1], respectively.

The human VPAC<sub>2</sub> receptor cDNA was cloned originally from a SUPT1 lymphoblastic cell line cDNA library [13]. We report here the characterisation of the structure of the human VIPR2 gene encoding the VPAC<sub>2</sub> receptor. The 5' and 3' untranslated regions (UTR) of the transcript have been determined by 5' and 3' rapid amplification of cDNA ends (RACE) of a human placental cDNA library with receptor specific primers. An additional 2.7 kb of the genomic 5' flanking region has been sequenced and a number of consensus binding sites for transcription factors which may be involved in the regulation of VIPR2 gene expression have been identified.

## 2. Materials and methods

### 2.1. Drugs and chemicals

Standard laboratory chemicals of Analar grade were obtained from Sigma or BDH Chemicals (Poole, UK). Oligonucleotide primers were obtained from Cruachem (Glasgow, UK), Oswell DNA Service (Edinburgh, UK) and Life Technologies (Paisley, UK).

### 2.2. Amplification of human VPAC<sub>2</sub> receptor from the human SUPT1 cell line

Total RNA was isolated using Catriomox-14 surfactant reagent (VH BIO, Newcastle-upon-Tyne, UK) from human SUPT1 cells (obtained from the European Collection of Animal Cell Culture, ECACC) which were grown in suspension. First strand cDNA was synthesised from 1 µg total RNA primed with random hexamers using SuperScript II reverse transcriptase (Life Technologies). Amplification of first strand cDNA for 35 cycles with the VPAC<sub>2</sub> receptor specific primers 6247 (5'-GGGGGACCTAGGACGGAG-3') and 4334 (5'-CCAGGTATGGGGTTAGTGGAC-3') and Pfu polymerase (Stratagene, Cambridge, UK) resulted in a 1.54 kb product which was gel-purified and ligated into the pGEM-T vector (Promega, Southampton, UK) and several clones were selected and sequenced in both directions.

### 2.3. Anchored 5' and 3' RACE for the human VPAC<sub>2</sub> receptor cDNA

An anchored cDNA library was synthesised from human placental poly-A<sup>+</sup> RNA using the Marathon cDNA Amplification kit (Clontech Laboratories, Basingstoke, Hampshire, UK). The 5' end of the human VPAC<sub>2</sub> receptor was amplified using touchdown PCR conditions and 35 cycles with the Clontech anchor primer, AP1, and the human VPAC<sub>2</sub> receptor specific primer 19076 (5'-AGGCAG-CAGCGTCCGCATCCCGAG-3'), with Advantage KlenTaq Polymerase Mix (Clontech). Amplification products were checked by gel electrophoresis. The first round of amplification reaction was then diluted 1:100 µl with sterile H<sub>2</sub>O and a second round of PCR was performed under the same conditions with 5 µl of diluted first round mix along with nested primers AP2 (Clontech) and the human VPAC<sub>2</sub> receptor specific primer EO5 (5'-GAGCGAGTGC GCGGAGACCT-

\*Corresponding author. Fax: (44) (131) 662 0240.  
E-mail: elutz@srv1.bmu.mrc.ac.uk

<sup>1</sup> Present address: University College London, Developmental Biology Unit, Institute of Child Health, 30 Guilford Street, London WC1N 1EH, UK.

3'). Amplification products were size-selected by gel electrophoresis, subcloned into the pGEM-T Easy vector (Promega) and sequenced in both directions. The 3' end of the VPAC<sub>2</sub> receptor was similarly amplified from anchored human placental cDNA, with primers AP1 and the human VPAC<sub>2</sub> receptor specific primer 32199 (5'-TGCCCTGGTGGGTCATACGAA-3') (for the first round of amplification) and with AP2 and the human VPAC<sub>2</sub> receptor specific primer 32200 (5'-AGCTGGTTGTCCACTAAACCCCAT-3') (for the second round). The 2.3 kb PCR product was gel-purified and subcloned into the pGEM-T Easy vector and several clones were selected and sequenced in both directions.

#### 2.4. Isolation of human genomic clones

A gridded human genomic DNA library RPCII (obtained from the UK MRC HGMP Resource Centre), constructed in the P1-PAC cloning vector pCYPAC2 [14], was screened with the full length human VPAC<sub>2</sub> receptor cDNA and with a 140 bp PCR fragment amplified from the cDNA with the exon 5 primer 9047 (5'-TCTGGTGAAGGCCATTTATACCC-3') and exon 6 primer HO4 (5'-CTCAGGATGAAGGACAGGAACAGG-3'). The cDNA probes were labelled with [ $\alpha$ -<sup>32</sup>P]dCTP (3300 mCi/mmol, NEN Life Science Products, Hounslow, UK) using Ready-To-Go DNA labelling beads (Amersham Pharmacia Biotech UK, Little Chalfont, Buckinghamshire, UK). Filters were hybridised at 45°C overnight in a buffer containing 50% formamide, 6×SSC (1×SSC=0.15 M NaCl, 0.015 M sodium citrate, pH 7.0), 5×Denhardt's solution, 0.5% SDS and 100 µg/ml sheared herring sperm DNA. After hybridisation, filters were washed at 65°C in 0.5×SSC, 0.5% SDS, 1 mM EDTA and exposed to Fuji RX autoradiography film. Positive clones were obtained from the UK MRC HGMP Resource Centre. Isolation of YAC clones HSC7E526 and E145A7 and cosmid clones 14, 126g6, 66e9, 122e11, 79g3 and 164g2 were as described by Mackay [1] and Shen et al. (in preparation).

Genomic *Bam*HI fragments were extracted from agarose gels with the QiaexII gel extraction kit (Qiagen, Crawley, UK), subcloned into pGEM11z (Promega) and sequenced with exon specific primers. Intron sizes were confirmed by PCR using human genomic DNA (Boehringer Mannheim, Lewes, UK), exon specific primers and the Expand Long Template PCR kit (Boehringer Mannheim) according to the manufacturer's instructions. Southern blots to detect the long YAC arm were as described by Shen et al. (in preparation).

### 3. Results

#### 3.1. Characterisation of the human VPAC<sub>2</sub> receptor cDNA

A single 1536 bp cDNA sequence encoding the VPAC<sub>2</sub> receptor was isolated from reverse transcriptase PCR amplification with the primers 6247 and 4334 of RNA extracted from the human lymphoblastic SUPT1 cell line (Fig. 1). The sequence of the cDNA contained an open reading frame (ORF) encoding a 438 amino acid protein which includes seven hydrophobic, putative transmembrane spanning domains. The 5' and 3' ends of the transcript were determined from 5' and 3' RACE fragments amplified from a human placental cDNA library. Thirty overlapping fragments of various lengths were isolated and sequenced after amplification of the 5' end. The sequence of the longest fragment extended 187 bp upstream of the ATG initiation codon. Amplification of the 3' end by RACE resulted in one 2.3 kb sequence. The 3' UTR extends 2473 bp downstream of the stop codon and includes a poly-adenylation signal sequence 19 residues upstream of the poly-A tail (Fig. 1). The composite cDNA sequence length is 3974 bp.

The pentamer sequence AUUUA is found within the VPAC<sub>2</sub> receptor cDNA sequence, along with U-rich domains and hairpin sequences. These have been implicated in the destabilisation and agonist-induced downregulation of receptor mRNA [15,16]. One AUUUA sequence is found in exon 5 (residues 385–389) and two in exon 13 in a region 190 bp downstream of the stop codon. The U-rich domains along with the hairpin sequence flank these sequences within the 3' UTR.

#### 3.2. The VIPR2 gene is comprised of 13 exons

In order to determine the structure of the VIPR2 gene, various genomic clones contained in YAC, PAC and cosmid vectors which hybridised with the VPAC<sub>2</sub> receptor cDNA were isolated. The YAC and cosmid clones have been described previously [1]. The PAC clone 118F4 was isolated from a gridded human genomic library after hybridisation with the 1.54 kb VPAC<sub>2</sub> receptor cDNA clone. A second PAC clone, 228M22, was isolated after rescreening the library with the 140 bp PCR fragment encoding exons 5 and 6 amplified from the human cDNA.

Intron sizes were determined by PCR amplification of human genomic DNA and Southern blot analysis of YAC, PAC and cosmid genomic clones (Table 1). Oligonucleotide primers were designed to amplify across putative exon boundaries which were predicted by characterisation of the mouse gene (Mackay et al., in preparation). Intron 4 was only partially characterised as it was too large, or unstable, for amplification and none of the PAC or cosmid clones spanned it. Intron 5 was amplified from human genomic DNA and PAC 228M22, but was too large (15 kb) to subclone for restriction and sequence analysis. Therefore, a 12 kb *Bam*HI fragment from PAC 228 DNA which hybridised with the exon 5/6 PCR fragment was excised from a gel and sequenced with exon 5 specific primers. Table 1 and Fig. 1 show the location of exon boundaries within the cDNA sequence.

Various *Bam*HI fragments from the PAC and cosmid clones, encoding exons 1, 2, 3-4, 6, 7-9 and 10-13, were subcloned and exon/intron boundaries were determined by sequencing. Fig. 2a shows the arrangement of exons along with a partial restriction map of the gene and position of the PAC and cosmid clones. The amino-terminus of the receptor is encoded by the first five exons, the transmembrane spanning domains by exons 5–12 and the carboxyl tail by exon 13 (Fig. 2b). To date, we have found no evidence for splice variants, or additional exons, for the VPAC<sub>2</sub> receptor.

#### 3.3. Human YAC clones containing the VIPR2 gene

The human VIPR2 gene was localised close to the telomere of chromosome 7q [1] with the 5' end of the gene oriented towards the telomere. The position of the VIPR2 gene and the presence of various chromosome 7q36 markers were determined for the YAC clones HSC7E526 and E145A7 by PCR analysis. Both the 550 kb YAC, HSC7E526, and the 200 kb YAC, E145A7, span the entire gene, which is 117 kb. Further characterisation of the YAC HSC7E526 by FISH suggested

Fig. 1. The 3974 bp sequence of the human VPAC<sub>2</sub> receptor (accession number X95097) and protein translation of the ORF. The transmembrane spanning domains are highlighted. Exon boundaries are defined by the symbol ▼. The arrows indicate the positions of the PCR primers used to amplify the cDNA (black arrows), the 5' RACE (stippled arrows) and 3' RACE (hatched arrows) fragments. The VIPR2 gene sequence has been deposited in the database, accession numbers Y18423–Y18431.

[illegible]

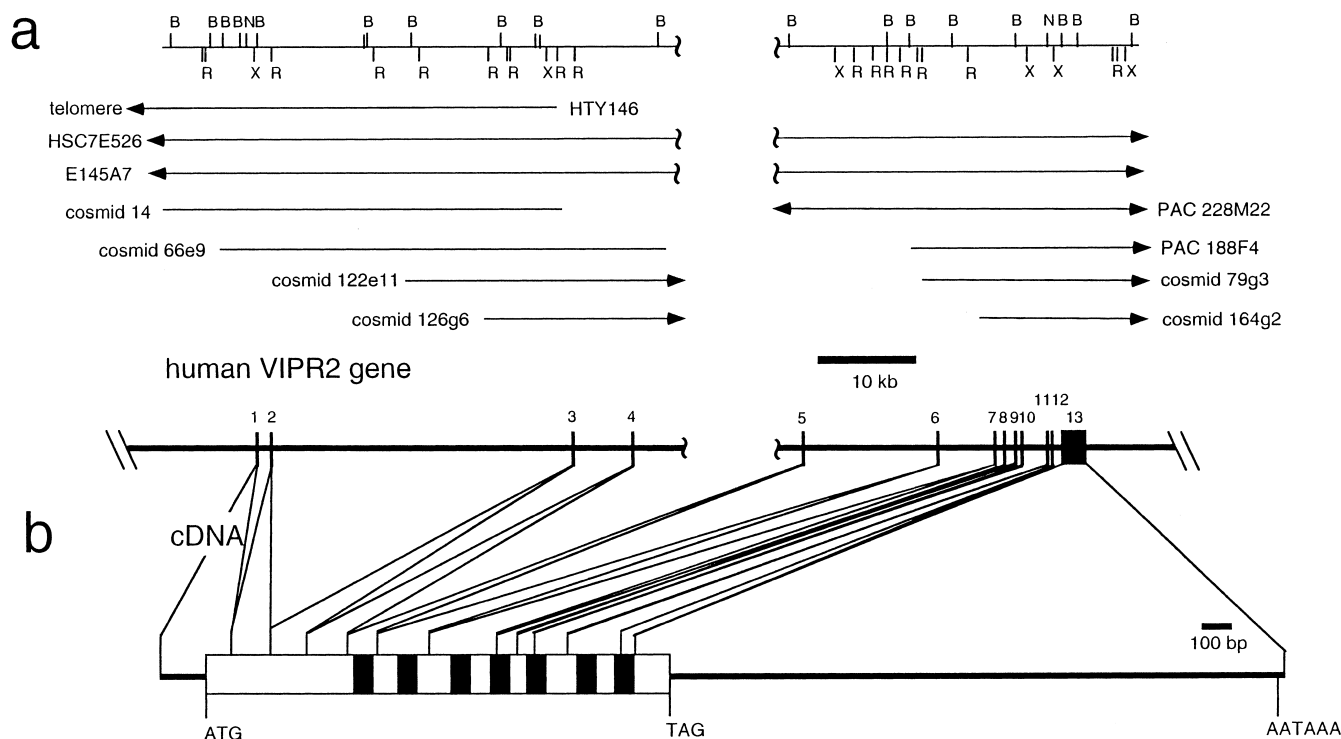


Fig. 2. (a) Partial restriction map of the human gene which includes sites for *EcoRI* (R), *XhoI* (X), *NotI* (N) and *BamHI* (B). The positions of YAC, PAC and cosmid clones are shown. Arrow heads indicate where genomic clones extend beyond the mapped region. Exons are indicated by the vertical lines and are numbered 1–13. The hash marks in intron 4 indicate the uncharacterised region. (b) Structure of the 3974 bp cDNA with the positions of exon boundaries indicated by lines descending from the exon regions marked in (a). The transmembrane spanning domains are indicated by the black boxes within the coding region, flanked by the ATG and TAG sequences. The position of the poly-adenylation signal is indicated by the AATAAA sequence.

that the YAC is not chimeric (data not shown). Analysis of the YAC HSC7E526 DNA with eight STS markers (D7S594, SWSS2567, SWSS1117, SWSS225, SWSS2253, D7S68, SWSS25 and SWSS3199) which span approximately 100 kb intervals showed that the YAC encompasses five (SWSS2567, SWSS1117, SWSS225, SWSS2253 and D7S68) of the markers tested (Fig. 3).

The complete sequence of the terminal 226 kb of human 7q telomere, including the entire VIPR2 promoter, exons 1 and 2, intron 1 and 31.9 kb of intron 2, was made available recently (accession number AF027390). This sequence indicates that the distance between the telomeric marker D7S594 and SWSS2567 is approximately 79 kb and the distance between D7S594 and SWSS1117 (which is located in intron 2 of the VIPR2 gene) is approximately 218 kb. Analysis of human

genomic DNA by PCR and by sequencing indicates that SWSS1117 lies approximately 0.6 kb upstream from exon 3 (data not shown). Subsequently, the entire gene sequence has been released (accession number AC007269), confirming the location of the marker SWSS225 within intron 4. The distance between exon 1 of the VIPR2 gene and the marker D7S594 is 184 kb. These figures suggest that the YAC HSC7E526 (truncated between markers D7S594 and SWSS2567) harbours ~105–180 kb of the 5' sequence, the entire VIPR2 gene (117 kb) and ~260–330 kb of the sequence downstream of the gene.

#### 3.4. The 5' promoter region is extremely GC-rich

The 5' promoter region was determined from the sequence of two overlapping genomic clones, a 2.4 kb *BamHI* and a 3.2

Table 1  
Table of exon/intron boundaries of the human VIPR2 gene

Exon boundaries	Splice junctions	Amino acid residue	Intron size (kb)
Exon 1/2	GCCCCGgtgagt...gaacagGTGAAC	Pro-17/Val-18	2.2
Exon 2/3	ACAAAGgtaagg...ttccagCCTGCA	Ala-51	32.5
Exon 3/4	AAGCAGgtactg...tttcagGAAACA	Gly-87	6.4
Exon 4/5	AGCAAGgtaggc...ttgttagATCACG	Lys-94/Ile-95	45
Exon 5/6	CTTCAGgtaagg...cctcagGAAGCT	Arg-152	15
Exon 6/7	TCCTGGgtaagg...ttccagGTGGGC	Trp-199/Val-200	6
Exon 7/8	GATGGGgtaagt...ctgcagGCCTCC	Gly-250	1.1
Exon 8/9	CACCGGgtgagt...tgacagTTGCTG	Gly-270	1.2
Exon 9/10	ATCATCgtaaag...gtgcagGTCAAT	Ile-293/Val-294	0.276
Exon 10/11	GTACAAgtgagt...gtgcagGAGGCT	Lys-324	2.1
Exon 11/12	TTCCAGgtaggt...ttgcagGGCCTG	Gln-367/Gly-368	0.068
Exon 12/13	AGTGAGgtaagt...tcacagGTGCAG	Glu-381/Val-382	1.1

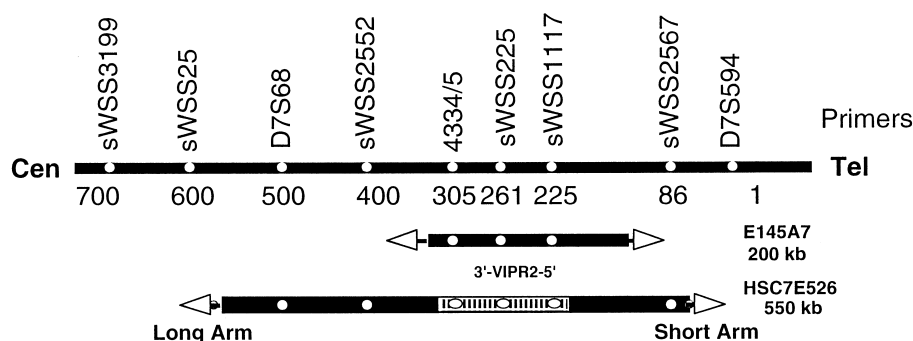


Fig. 3. Physical map of the terminal region of human chromosome 7q36 and position of the VIPR2 gene contained within YAC clones HSC7E526 and E145A7. The order of the DNA markers at the extreme end of 7q36 and their location relative to the gene are shown.

kb *Pst*I fragment subcloned from cosmid 14, which included the ATG initiation codon (exon 1), the 5' flanking sequence and part of intron 1. Identification of transcription factor sites was performed using the TESS-String-based Search computer program<sup>2</sup>. Fig. 4 shows part of the 5' flanking sequence, along with the start of the longest 5' RACE product (used to define the 5' boundary of exon 1) and location of transcription factor sites which may be involved in regulating expression of the VIPR2 gene. Primer extension was tried to determine the exact location of the transcription start site, but was not successful perhaps due to the high GC content of the 5' UTR (80%). The GC content of the 5' flanking sequence extending 850 bp upstream of the 5' RACE product is approximately 71%.

No TATA-box was found in the GC-rich region upstream of the putative 5' boundary of exon 1. Within the GC-rich region are several sites for the transcription factor Sp1 [17]. Immediately upstream of the GC-rich region is a 1.35 kb region which contains multiple repeats of the sequences CAGATAG (14 repeats) and GATGATAG (25 repeats) which are potential sites for the GATA family of transcription factors [18].

#### 4. Discussion

In this study, we have characterised the gene structure of the human VPAC<sub>2</sub> receptor and defined the 5' and 3' ends of the cDNA. The 3974 bp cDNA sequence was determined from a composite which included the sequences of the 5' and 3' RACE products. The size of the human transcript has previously been determined by Northern blot analysis as two hybridising bands at 4.6 and 2.3 kb [19]. The presence of AUUUA pentamer sequences and U-rich regions within the coding sequence and 3' UTR provides evidence for post-transcriptional regulation of the mRNA [20]. These regions have been shown to be important for protein binding to the 5-HT<sub>2c</sub> receptor mRNA [16] and are involved in agonist induced destabilisation of the  $\beta$ -adrenergic receptor mRNA [15]. Similar regions are found within the calcitonin receptor mRNA [21].

We were unable to determine the exact start of transcription for the VIPR2 gene. Some members of the secretin receptor

family have multiple exons encoding the 5' UTR [21–24]. However, this does not appear to be the case for the VPAC<sub>2</sub> receptor. The 5' UTR (determined from the 5' RACE product) and the translational start site of the human gene are encoded within the same exon. As with the mouse *Vipr2* gene (Mackay et al., in preparation), this exon and the genomic sequence flanking it are extremely GC-rich. The human exon is located within a 1.4 kb CpG island. Although there are no TATA-box sequences near the 5' end of this exon, there are several consensus binding sites for the transcription factors Sp1 and MAZ, which have been shown to be important for transcription of the 5-HT<sub>1a</sub> receptor gene which has a TATA-less promoter region [25]. Promoter regions which are GC-rich and contain Sp1 binding sites but not TATA sites have been shown to be activated by Sp1 recruiting specific co-activators to bind TFIID [26,27]. The promoter regions which have been characterised for genes of members of the secretin receptor family typically do not have a TATA-box immediately upstream of the transcriptional start site and are extremely GC-rich [12,21,28–30]. Pei et al. [30] have shown that Sp1 binds to the promoter and activates the rat *Vipr1* gene. We are currently investigating whether a similar transcription mechanism, involving Sp1 rather than TATA binding proteins, operates for the VIPR2 gene.

Multiple promoter regions have been described for PTH and glucagon receptor genes. The PTH receptor gene has been shown to be controlled by two promoters, one of which is not GC-rich and is only active in kidney [31]. There is no evidence that the VPAC<sub>2</sub> receptor is controlled by more than one promoter as only one type of 5' UTR has been found with 5' RACE.

A number of putative binding sites for tissue specific factors are located within the genomic 5' flanking sequence. For instance, a site for the transcription factor Pit-1A, which is involved in activating gene expression in specific pituitary cell types [32], is found approximately 1 kb upstream of exon 1. The VPAC<sub>2</sub> receptor is expressed in the anterior pituitary and in clonal somatomammotroph cell lines [33]. Also within this region are sites for the CAC binding protein [34] and LyF-1 [35], further upstream of the VIPR2 exon 1 is a region (between residues –1200 and –2550) which contains multiple repeats and GATA sites [18]. These sites may be involved in the differential expression of the VPAC<sub>2</sub> receptor in lymphocytes [36].

Genes encoding receptors belonging to the secretin receptor family are interrupted by several introns [12,21–23,31,37]. The VIPR2 gene is composed of 13 exons, which are interrupted

<sup>2</sup> TESS: Transcription Element Search System on the World Wide Web; Jonathan Schug and G. Christian Overton, Technical Report CBIL-TR-1997-1001-v0.0 of the Computational Biology and Informatics Laboratory, School of Medicine, University of Pennsylvania, 1997.

-2600 TTTCTTCACAAGCCAAGCAGCCCGCCTAAGGCCAGGTCTCCCTCTCTCGTCTCCCACTAGATACACATGGATGAAGCCAGATAGATAACAGGTAGAGG  
 -2500 ATAGATGACAGGCAGATGACAGATAGGCAGATGATAGATGATCGATAGATGACAGGCAGATGATGGATGACAGGCAGATCACAGATATGATAGATCACAG  
 -2400 ATAGATGACAAATAGGCAGATGATAGATGACAGGCAGATGACAGATAGGCAGATGATAGATGACTGATAGATGACAGGCAGATCACAGATATGATAGATC  
 -2300 ACAGATAGATGACAAATAGGCAGATGATAGATGACAGGCAGATGACAGATAGGCAGATGATAGATGACTGATAGATGACAGGCAGATCACAGATATGATA  
 -2200 GATCACAGATAGATGACAAATAGGCAGATGATAGATGACTGATAGATGACAGGCAGATCACAGATATGATAGATCACAGATAGATGATAAATAGGCAGAT  
 -2100 GATAGATGACAGGCAGATGACAGATAGGCAGATGATAGATGATAGATGACAGGCAGATGACAGATAGGCAGATGATAGATTGATAGATGACAGGCAGAT  
 -2000 GACAGATAGGCAGATGATAGGTGATTGATAGATGACAGGCAGATGACAGATAGGCAGATGATAGGTGATTGATAGATGACAGGCAGGTAGATCACAGATA  
 -1900 TGACGGATCACAGATGATGGATAGGCAGATGATAGATGACAGATAATAGATGATAGATTACAGATAGACGATCGATAGGAAGATAATAGATGATAGGTAC  
 -1800 ATGACAGATCATAGATAAATGACAGAGAGGTAGATGATAGATGATAAATAGGTAGATGGTAGGTAAATAGATAGGTAAATAGATCGCAGAGAGATGATAGGT  
 -1700 AGATGATAGCTGATTGATACATGATACAGAGATGATAGGTACATGATAGATGGATAGATCATAAATGACAGCAGATGATAGATGATAAATAGGTAGATGA  
 -1600 CAGGTATAGATAGGTAAATTCAGATGAAAGATAGGCAGATGATAGATGATAGGTACATAATAGATGGATAGATCATAGATAATAGATGACAGGTAGGT  
 -1500 GATAGAGGAAAGGTGATAGATGATAGGTAGGCAGATGATAGATGATAGGTACATGACAGATGGATAGATGAAAGGTAGATGATGGACAGGTAGATGATAGATGA  
 -1400 CAGATAATAGATGACAGATTACAGATAGATGATCGATAGGAAGATGATAGATGATAGGTACATGACAGATCGATAGATGATACGTAGATGATGGATAGGT  
 -1300 AGATGATAGATAATGTGGATGATAAATCAGATGATAGAAAAATCGATATCTGTGAATACGGGGGTGGGGATCAGAGAGAAGCAAAATGATAAGCCCA  
 CAC-binding protein CAAT  
 -1200 ATGAGGTAAATGTTAATGACCGAATCCACAAAAAGGATAAAAAGGAGTTATTTGTAATAATCTTATTTCTTAAATTTTGTGTAAGTTTGAATTAATTTT  
 Pitla C/EBP  
 -1100 CAAATAAAAAATTAATAACACGAACCTTAACCTTGTATCTCTCATTTCCAGTGTCTTCCACTTGCTGGGGAACGCCGAGCTCTCTGGGTGGTCACGCCG  
 LyF-1 (SP1) CAC-binding protein  
 -1000 GCGCCTTGGGAGGCCGCGCAGTCCCGAGTGGGGCGCGCGGGCGGCCCGCAGTGGCAGCCGCTGAGTCCCCCGCCAGCGTTCCCCACCCGCGCGCGTT  
 -900 TCGGGGAGAGAATGACCCCGTTTTCGAAACGCAGGACACAAAACCCGCCACCCAGCTGAGGCTGGGACCCACCTATACCCCGTCTGCGTTTTCCTCC  
 NF-κB, CAC-binding protein  
 -800 AGCTTTCTGTTCGACCTGATGCTCTTTGAGGCTCGGGGATCCCCACCAAAAGCACTCTGATTTTCTCCATTTTCAGATGCCCCAACCTAGCCCCAC  
 -700 TGGCTACAAAGACGGCGCGCCACCCCGGACGGGTCTCTGGCCGTTTCTTGGGCTTCCGAGTTGTTTGAAGAGCAGGTACAGGCGCGCGCAAGGCA  
 -600 TCCGCCCCCTTCGAGCGCACACCTGGCCCGTCTCTCTCCCGCCCCGCGCGCGCGCTCTCCCCAGTTCTCTACTCGGAGCGGGCGCGCACTG  
 MAZ/SP1 (TFIID)  
 -500 AGCACCCCGTTCCAGCCCTCCGCCCACAGATCGGGGAGGCGCGCGGGAGGGGGCGTGGGCCACGGAAGAGGGGAGGAGAACGGAGAGAGAGGCTG  
 IF-1 SP1  
 -400 GGGGGTTCGAGAGACAGAGGAGCGCGCGCGGGGACAGAGAGCCAGGGGCGAGGAGGGCGCGGGCGCAGCGGAAGGGGAAGTGGGGGCGG  
 (C/EBP) (C/EBP) (CAAT) SP1 SP1  
 -300 GGAGGAGGGGCGCGGAGCGGGGTTCTCGGGGAGGAGGAGACGAGGATTGGGGCAGCGCGGGCGGGGACAGGGGCGGGGCGGAGCGGCGGGCGG  
 ⇒ ▼  
 -200 GTGGGGACGGGCGTGCATTGAGCGCGCTCCAGCTGCCGGGACGGAGGGGGCGCCCCCGCTCGGGGCGCTCGGCTACAGCTGCGGGGCCGAGGTCT  
 ▽  
 -100 CCGCGCACTCGCTCCCGGCCATGCTGGAGGCGCGGAACCGGGGACCTAGGACGAGGCGGGCGCGCTGGGCGCCCCCGCACGCTGAGCTCGGG  
 +1 ATG

Fig. 4. Sequence of the 5' region flanking the translation start site within the first exon of the VIPR2 gene. The translation start site (in bold) is numbered +1 and the sequence comprising the first exon is italicised. The symbol ⇒ marks the start of the longest 5' RACE product and the boundary of exon 1. The position of the 3' end of the primer 19076 is marked with the symbol ▽ and the primer EO5 with the symbol ▼. Putative transcription factor binding sites in the promoter region are labelled above the sequence. The GATA site sequences are underlined.

by introns of varying length from 68 bp (intron 11) to 45 kb (intron 4). This gene appears to be one of the largest genes in this receptor family, spanning 117 kb. Other receptors belonging to this family are alternatively spliced to generate various isoforms of the receptor. The rat PAC<sub>1</sub> receptor can be alternatively spliced in the 5' UTR [24], the amino-terminal extracellular domain [38] and the third intracellular loop (ic3) [39]. The region encoding ic3 is interrupted by an intron which contains two additional 84 bp exons (hip and hop) which give rise to one of two 28 amino acid cassettes (hip and hop1), a 27 amino acid cassette (hop2), a 56 amino acid cassette (hiphop1) or a 55 amino acid cassette (hiphop2) which may be spliced into ic3 at amino acid residue L-348, giving rise to six iso-

forms of this receptor [39]. Similarly, the two exons encoding the GHRH receptor ic3 are interrupted by an intron which contains a 123 bp exon encoding a 41 amino acid cassette [40]. The long form of the GHRH receptor has this cassette inserted in ic3 at the same position as those of the PAC<sub>1</sub> receptor. The human VPAC<sub>2</sub> receptor ic3 is encoded by exons 10 and 11, which are interrupted by a 2.1 kb intron. We have sequenced this intron (accession number Y18430) and have found a number of consensus sequences conforming to the GT-AG rule for exon/intron splice junctions [41]. However, amplification of this region from cDNA made from various tissues has not indicated the presence of other isoforms of the VPAC<sub>2</sub> receptor.

In summary, we have described the 5' and 3' ends of the human VPAC<sub>2</sub> receptor cDNA and characterised the exon/intron organisation of the VIPR2 gene. In addition, we have identified several putative regulatory sites within a GC-rich sequence flanking the 5' end of the first exon.

**Acknowledgements:** We thank Rory Mitchell and Melanie Johnson for critical discussions and helpful advice, Derek Robertson and Elma Clark for technical assistance and Marianne Eastwood for help in preparing this manuscript. Eve Lutz is a member of the Membrane Biology Interdisciplinary Research Group of the University of Edinburgh.

## References

- [1] Mackay, M., Fantes, J., Scherer, S., Boyle, S., Lutz, E., Van Heyningen, V. and Harmar, A.J. (1996) *Genomics* 37, 345–353.
- [2] Segre, V. and Goldring, S.R. (1993) *Trends Endocrinol. Metab.* 4, 309–314.
- [3] Ishihara, T., Shigemoto, R., Mori, K., Takahashi, K. and Nagata, S. (1992) *Neuron* 8, 811–819.
- [4] Lutz, E.M., Sheward, W.J., West, K.M., Morrow, J.A., Fink, G. and Harmar, A.J. (1993) *FEBS Lett.* 334, 3–8.
- [5] Harmar, A.J., Arimura, A., Gozes, I., Journot, L., Laburthe, M., Pisegna, J.R., Rawlings, S.R., Robberecht, P., Said, S.I., Sreedharan, S.P., Wank, S.A. and Waschek, J.A. (1998) *Pharmacol. Rev.* 50, 265–270.
- [6] Ishihara, T., Nakamura, S., Kaziyo, Y., Takahashi, T., Takahashi, K. and Nagata, S. (1991) *EMBO J.* 10, 1635–1641.
- [7] Pisegna, J. and Wank, S. (1993) *Proc. Natl. Acad. Sci. USA* 90, 6345–6349.
- [8] Morrow, J.A., Lutz, E.M., West, K.M., Fink, G. and Harmar, A.J. (1993) *FEBS Lett.* 329, 99–105.
- [9] Usdin, T.B., Bonner, T.I. and Mezey, E. (1994) *Endocrinology* 135, 2662–2680.
- [10] Sheward, W.J., Lutz, E.M. and Harmar, A.J. (1995) *Neuroscience* 67, 409–418.
- [11] Inagaki, N., Yoshida, H., Mizuta, M., Mizuno, N., Fujii, Y., Gono, T., Miyazaki, J.-I. and Seino, S. (1994) *Proc. Natl. Acad. Sci. USA* 91, 2679–2683.
- [12] Sreedharan, S.P., Huang, J.-X., Cheung, M.-C. and Goetzl, E.J. (1995) *Proc. Natl. Acad. Sci. USA* 92, 2939–2943.
- [13] Svoboda, M., Tastenoy, M., Van Rampelbergh, J., Goossens, J.-F., De Neef, P., Waelbroeck, M. and Robberecht, P. (1994) *Biochem. Biophys. Res. Commun.* 205, 1617–1624.
- [14] Ioannou, P.A. and de Jong, P.J. (1996) in: *Construction of Bacterial Artificial Chromosome Libraries Using the Modified P1 (PAC) System. Current Protocols in Human Genetics*, John Wiley and Sons, New York.
- [15] Huang, L.-Y., Tholanikunnel, B.G., Vakalopoulou, E. and Malbon, C.C. (1993) *J. Biol. Chem.* 268, 25769–25775.
- [16] Kao, H.-T., Ghafoori, S., Porton, B., Wong, D.L. and Ciaranello, R.D. (1996) *Mol. Brain Res.* 43, 174–184.
- [17] Kadonaga, J.T., Jones, K.A. and Tjian, R. (1986) *Trends Biochem. Sci.* 11, 20–23.
- [18] Whyatt, D.J., de Boer, E. and Grosveld, F. (1993) *EMBO J.* 12, 4993–5005.
- [19] Adamou, J.E., Aiyar, N., Van Horn, S. and Elshourbagy, N.A. (1995) *Biochem. Biophys. Res. Commun.* 209, 385–392.
- [20] Malter, J.S. (1989) *Science* 246, 664–666.
- [21] Zolnierowicz, S., Cron, P., Solinas-Toldo, S., Fries, R., Lin, H.Y. and Hemmings, B.A. (1994) *J. Biol. Chem.* 269, 19530–19538.
- [22] McCuaig, K.A., Clarke, J.C. and White, J.H. (1994) *Proc. Natl. Acad. Sci. USA* 91, 5051–5055.
- [23] Aino, H., Hashimoto, H., Ogawa, N., Nishino, A., Yamamoto, K., Nogi, H., Nagata, S. and Baba, A. (1995) *Gene* 164, 301–304.
- [24] Chatterjee, T.K., Liu, X., Davisson, R.L. and Fisher, R.A. (1997) *J. Biol. Chem.* 272, 12122–12131.
- [25] Parks, C.L. and Shenk, T. (1996) *J. Biol. Chem.* 271, 4417–4430.
- [26] Pugh, B.F. and Tjian, R. (1990) *Cell* 61, 1187–1197.
- [27] Courey, A.J., Holtzman, D.A., Jackson, S.P. and Tjian, R. (1989) *Cell* 59, 827–836.
- [28] Burcelin, R., Li, J. and Charron, M.J. (1995) *Gene* 164, 305–310.
- [29] Lankat-Buttgereit, B. and Göke, B. (1997) *Peptides* 18, 617–624.
- [30] Pei, L. and Melmed, S. (1995) *Biochem. J.* 308, 719–723.
- [31] McCuaig, K.A., Lee, H.S., Clarke, J.C., Assar, H., Horsford, J. and White, J.H. (1995) *Nucleic Acids Res.* 23, 1948–1955.
- [32] Nelson, C., Albert, V.R., Elsholtz, H.P., Lu, L.I.-W. and Rosenfeld, M.G. (1988) *Science* 239, 1400–1405.
- [33] Rawlings, S.R., Piuze, I., Schlegel, W., Bockaert, J. and Journot, L. (1995) *Endocrinology* 136, 2088–2098.
- [34] Wang, Y., Kobori, J.A. and Hood, L. (1993) *Mol. Cell. Biol.* 13, 5691–5701.
- [35] Lo, K., Landau, N.R. and Smale, S.T. (1991) *Mol. Cell. Biol.* 11, 5229–5243.
- [36] Delgado, M., Martinez, C., Johnson, M.C., Gonariz, R.P. and Ganea, D. (1996) *J. Neuroimmunol.* 68, 27–38.
- [37] Lok, S., Kuijper, J.L., Jelinek, L.J., Kramer, J.M., Whitmore, T.E., Sprecher, C.A., Mathewes, S., Grant, F.J., Biggs, S.H., Rosenberg, G.B., Sheppard, P.O., O'Hara, P.J., Foster, D.C. and Kindsvogel, W. (1994) *Gene* 140, 203–209.
- [38] Pantaloni, C., Brabet, P., Bilanges, B., Dumuis, A., Houssami, S., Spengler, D., Bockaert, J. and Journot, L. (1996) *J. Biol. Chem.* 271, 22146–22151.
- [39] Spengler, D., Waeber, C., Pantaloni, C., Holsboer, F., Bockaert, J., Seeburg, P.H. and Journot, L. (1993) *Nature* 365, 170–175.
- [40] Mayo, K.E. (1992) *Mol. Endocrinol.* 6, 1734–1744.
- [41] Breathnach, R., Benoist, C., O'Hare, K., Gannon, F. and Chambon, P. (1978) *Proc. Natl. Acad. Sci. USA* 75, 4853–4857.